

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

TAG 675

- (2) INFORMATION FOR SEQUENCE ID NO: 3:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 228 base pairs
 (B) TYPE: nucleic acid
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: genomic DNA
(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GCATGCAGTT	GCAAAGCCCA	GAAGAAAGAA	ATGGACAGCG	GAAGAAGTGG	TTGTTTTTTT	60
TTCCCCTTCA	TTAATTTTCT	AGTTTTTAGT	AATCCAGAAA	ATTTGATTTT	GTTCTAAAGT	120
TCATTATGCA	AAGATGTCAC	CAACAGACTT	CTGACTGCAT	GGTGAACTTT	CATATGATAC	180
ATAGGATTAC	ACTTGTACCT	GTAAAAATA	AAAGTTTGAC	TTGCATAC		228

(2) INFORMATION FOR SEQUENCE ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1365 base pairs

(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ACCACAGGAG	AATGAAAAGA	ACCCGGGACT	CCCAAAGACG	CTAGATGTGT	50
GAAGATCCTG	ATCACTCATT	GGGTGTCTGA	GTTCTGCGAT	ATTCATCCCT	100
CAGCCAATGA	GCTTACTGTT	CTCGTGGGGG	GTTTGTGAGC	CTTGGGTAGG	150
AAGTTTGTCA	AGTTCCGCCT	ACAGCTCTAG	CTTGTGAATT	TGTACCCTTT	200
CACGTAAAAA	AGTAGTCCAG	AGTTTACTAC	ACCCTCCCTC	CCCCCTCCCA	250
CCTCGTGCTG	TGCTGAGTTT	AGAAGTCTTC	CTTATAGAAG	TCTTCCGTAT	300
AGAACTCTTC	CGGAGGAAGG	AGGGAGGACC	CCCCCCTTTT	GCTCTCCCAG	350
CATGCATTGT	GTCAACGCCA	TTGCACTGAG	CTGGTCGAAG	AAGTAAGCCG	400
CTAGCTTGCG	ACTCTACTCT	TATCTTAACT	TAGCTCGGCT	TCCTGCTGGT	450
ACCCTTTGTG	CC				462
ATG TCT GAT	AAC AAG AAA CCA GAC	AAA GCC CAC	AGT GGC TCA		504
GGT GGT GAC	GGT GAT GGG AAT AGG	TGC AAT TTA	TTG CAC CGG		546
TAC TCC CTG	GAA GAA ATT CTG CCT	TAT CTA GGG	TGG CTG GTC		588
TTC GCT GTT	GTC ACA ACA AGT TTT	CTG GCG CTC	CAG ATG TTC		630
ATA GAC GCC	CTT TAT GAG GAG CAG	TAT GAA AGG	GAT GTG GCC		672
TGG ATA GCC	AGG CAA AGC AAG CGC	ATG TCC TCT	GTC GAT GAG		714
GAT GAA GAC	GAT GAG GAT GAT GAG	GAT GAC TAC	TAC GAC GAC		756
GAG GAC GAC	GAC GAC GAT GCC TTC	TAT GAT GAT	GAG GAT GAT		798
GAG GAA GAA	GAA TTG GAG AAC CTG	ATG GAT GAT	GAA TCA GAA		840
GAT GAG GCC	GAA GAA GAG ATG AGC	GTG GAA ATG	GGT GCC GGA		882
GCT GAG GAA	ATG GGT GCT GGC GCT	AAC TGT GCC	TGT GTT CCT		924
GGC CAT CAT	TTA AGG AAG AAT GAA	GTG AAG TGT	AGG ATG ATT		966
TAT TTC TTC	CAC GAC CCT AAT TTC	CTG GTG TCT	ATA CCA GTG		1008
AAC CCT AAG	GAA CAA ATG GAG TGT	AGG TGT GAA	AAT GCT GAT		1050
GAA GAG GTT	GCA ATG GAA GAG GAA	GAA GAA GAA	GAG GAG GAG		1092
GAG GAG GAA	GAG GAA ATG GGA AAC	CCG GAT GGC	TTC TCA CCT		1134
TAG					1137
GCATGCAGTT	GCAAAGCCCA	GAAGAAAGAA	ATGGACAGCG	GAAGAAGTGG	1187
TTGTTTTTTT	TTCCCCTTCA	TTAATTTTCT	AGTTTTTAGT	AATCCAGAAA	1237
ATTTGATTTT	GTTCTAAAGT	TCATTATGCA	AAGATGTCAC	CAACAGACTT	1287
CTGACTGCAT	GGTGAACCTT	CATATGATAC	ATAGGATTAC	ACTTGTACCT	1337
GTTAAAAATA	AAAGTTTGAC	TTGCATAC			1365

(2) INFORMATION FOR SEQUENCE ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4698 base pairs

(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ACCACAGGAG	AATGAAAAGA	ACCCGGGACT	CCCAAAGACG	CTAGATGTGT	50
GAAGATCCTG	ATCACTCATT	GGGTGTCTGA	GTTCTGCGAT	ATTCATCCCT	100
CAGCCAATGA	GCTTACTGTT	CTCGTGGGGG	GTTTGTGAGC	CTTGGGTAGG	150
AAGTTTGTCA	AGTTCCGCCT	ACAGCTCTAG	CTTGTGAATT	TGTACCCTTT	200
CACGTAAAAA	AGTAGTCCAG	AGTTTACTAC	ACCCTCCCTC	CCCCCTCCCA	250
CCTCGTGCTG	TGCTGAGTTT	AGAAGTCTTC	CTTATAGAAG	TCTTCCGTAT	300
AGAACTCTTC	CGGAGGAAGG	AGGGAGGACC	CCCCCCTTTT	GCTCTCCCAG	350
CATGCATTGT	GTCAACGCCA	TTGCACTGAG	CTGGTCGAAG	AAGTAAGCCG	400
CTAGCTTGCG	ACTCTACTCT	TATCTTAACT	TAGCTCGGCT	TCCTGCTGGT	450
ACCCTTTGTG	CC				462
ATG TCT GAT	AAC AAG AAA	CCA GAC AAA	GCC CAC AGT	GGC TCA	504
GGT GGT GAC	GGT GAT GGG	AAT AGG TGC	AAT TTA TTG	CAC CGG	546
TAC TCC CTG	GAA GAA ATT	CTG CCT TAT	CTA GGG TGG	CTG GTC	588
TTC GCT GTT	GTC ACA ACA	AGT TTT CTG	GCG CTC CAG	ATG TTC	630
ATA GAC GCC	CTT TAT GAG	GAG CAG TAT	GAA AGG GAT	GTG GCC	672
TGG ATA GCC	AGG CAA AGC	AAG CGC ATG	TCC TCT GTC	GAT GAG	714
GAT GAA GAC	GAT GAG GAT	GAT GAG GAT	GAC TAC TAC	GAC GAC	756
GAG GAC GAC	GAC GAC GAT	GCC TTC TAT	GAT GAT GAG	GAT GAT	798
GAG GAA GAA	GAA TTG GAG	AAC CTG ATG	GAT GAT GAA	TCA GAA	840
GAT GAG GCC	GAA GAA GAG	ATG AGC GTG	GAA ATG GGT	GCC GGA	882
GCT GAG GAA	ATG GGT GCT	GGC GCT AAC	TGT GCC T		916
GTGAGTAACC	CGTGGTCTTT	ACTCTAGATT	CAGGTGGGGT	GCATTCTTTA	966
CTCTTGCCCA	CATCTGTAGT	AAAGACCACA	TTTTGGTTGG	GGGTCATTGC	1016
TGGAGCCATT	CCTGGCTCTC	CTGTCCACGC	CTATCCCCGC	TCCTCCCATC	1066
CCCCACTCCT	TGCTCCGCTC	TCTTTCCTTT	TCCCACTTGG	CCTCTGGAGC	1116
TTCAGTCCAT	CCTGCTCTGC	TCCCTTTCCC	CTTTGCTCTC	CTTGCTCCCC	1166
TCCCCCTCGG	CTCAACTTTT	CGTGCCTTCT	GCTCTCTGAT	CCCCACCCTC	1216
TTCAGGCTTC	CCCATTTGCT	CCTCTCCCGA	AACCCTCCCC	TTCCTGTTCC	1266
CCTTTTCGCG	CCTTTTCTTT	CCTGCTCCCC	TCCCCCTCCC	TATTTACCTT	1316
TCACCAGCTT	TGCTCTCCCT	GCTCCCCCTC	CCCTTTTGCA	CCTTTTCTTT	1366
TCCTGCTCCC	CTCCCCCTCC	CCTCCCTGTT	TACCCTTCAC	CGCTTTTCCT	1416
CTACCTGCTT	CCCTCCCCCT	TGCTGCTCCC	TCCCTATTTG	CATTTTCGGG	1466
TGCTCCTCCC	TCCCCCTCCC	CCTCCCTCCC	TATTTGCATT	TTCGGGTGCT	1516
CCTCCCTCCC	CCTCCCCAGG	CCTTTTTTTT	TTTTTTTTTT	TTTTTTTTTT	1566
TTGGTTTTTC	GAGACAGGGT	TTCTCTTTGT	ATCCCTGGCT	GTCCTGGCAC	1616
TCACTCTGTA	GACCAGGCTG	GCCTCAAAC	CAGAAATCTG	CCTGCCTCTG	1666
CCTCCCAAAT	GCTGGGATTA	AAGGCTTGCA	CCAGGACTGC	CCCAGTGCAG	1716
GCCTTTCCTT	TTTCTCCTCT	CTGGTCTCCC	TAATCCCTTT	TCTGCATGTT	1766
AACTCCCCTT	TTGGCACCTT	TCCTTTACAG	GACCCCTCC	CCCTCCCTGT	1816
TTCCCTTCCG	GCACCCTTCC	TAGCCCTGCT	CTGTTCCCTC	TCCCTGCTCC	1866
CCTCCCCCTC	TTTGCTCGAC	TTTGTAGCAG	CTTACCTCTC	CCTGCTTTCT	1916
GGCCCGTTCC	CCTTTTTTGT	GCCTTTCCTC	CTGGCTCCCC	TCCACCTTCC	1966

AGCTCACCTT	TTTGTTTGTT	TGGTTGTTTG	GTTGTTTGGT	TTGCTTTTTT	2016
TTTTTTTTTT	GCACCTTGTT	TTCCAAGATC	CCCCTCCCCC	TCCGGCTTCC	2066
CCTCTGTGTG	CCTTTCCTGT	TCCCTCCCCC	TCGCTGGCTC	CCCCTCCCTT	2116
TCTGCCTTTC	CTGTCCCTGC	TCCCTTCTCT	GCTAACCTTT	TAATGCCTTT	2166
CTTTTCTAGA	CTCCCCCCTC	CAGGCTTGCT	GTTTGCTTCT	GTGCACTTTT	2216
CCTGACCCTG	CTCCCCCTTC	CCTCCCAGCT	CCCCCTCTT	TTCCCACCTC	2266
CCTTTCTCCA	GCCTGTCACC	CCTCCTTCTC	TCCTCTCTGT	TTCTCCCACT	2316
TCCTGCTTCC	TTTACCCCTT	CCCTCTCCCT	ACTCTCCTCC	CTGCCTGCTG	2366
GACTTCCTCT	CCAGCCGCCC	AGTTCCTGCT	AGTCCTGGAG	TCTTTCCTGC	2416
CTCTCTGTCC	ATCACTTCCC	CCTAGTTTCA	CTTCCCTTTC	ACTCTCCCCT	2466
ATGTGTCTCT	CTTCCCTATCT	ATCCCTTCCT	TTCTGTCCCC	TCTCCTCTGT	2516
CCATCACCTC	TCTCCTCCCT	TCCCTTTCCT	CTCTCTTCCA	TTTTCTTCCA	2566
CCTGCTTCTT	TACCCTGCCT	CTCCCATTCG	CCTCTTACCT	TTATGCCCCAT	2616
TCCATGTCCC	CTCTCAATTC	CCTGTCCCAT	TGTGCTCCCT	CACATCTTCC	2666
ATTTCCCTCT	TTCTCCCTTA	GCCTCTTCTT	CCTCTTCTCT	TGTATCTCCC	2716
TTCCCTTTGC	TTCTCCCTCC	TCCTTTCCCC	TTCCCCTATG	CCCTCTACTC	2766
TACTTGATCT	TCTCTCCTCT	CCACATAACC	TTTTTCCTTT	CCACCCTGCC	2816
CTTTGTCCCC	AGACCCTACA	GTATCCTGTG	CACAGGAAGT	GGGAGGTGCC	2866
ATCAACAACA	AGGAGGCAAG	AAACAGAGCA	AAATCCCCAA	ATCAGCAGGA	2916
AAGGCTGGAT	GAAAATAAGG	CCAGGTTCTG	AGGACAGCTG	GAATCTAGCC	2966
AAGTGGCTCC	TATAACCCTA	AGTACCAAGG	GAGAAAGTGA	TGGTGAAGTT	3016
CTTGATCCTT	GCTGCTTCTT	TTACATATGT	TGGCACATCT	TTCTCAAATG	3066
CAGGCCATGC	TCCATGCTTG	GCGCTTGCTC	AGCGTGGTTA	AGTAATGGGA	3116
GAATCTGAAA	ACTAGGGGCC	AGTGGTTTGT	TTTGGGGACA	AATTAGCACG	3166
TAGTGATATT	TCCCCCTAAA	AATTATAACA	AACAGATTCA	TGATTTGAGA	3216
TCCTTCTACA	GGTGAGAAGT	GGAAAAATTG	TCACTATGAA	GTTCTTTTTA	3266
GGCTAAAGAT	ACTTGGAACC	ATAGAAGCGT	TGTTAAAATA	CTGCTTTCTT	3316
TTGCTAAAAT	ATTCTTTCTC	ACATATTCAT	ATTCTCCAG		3355
GT GTT CCT	GGC CAT CAT	TTA AGG	AAG AAT GAA	GTG AAG TGT	3396
AGG ATG ATT	TAT TTC TTC	CAC GAC	CCT AAT TTC	CTG GTG TCT	3438
ATA CCA GTG	AAC CCT AAG	GAA CAA	ATG GAG TGT	AGG TGT GAA	3480
AAT GCT GAT	GAA GAG GTT	GCA ATG	GAA GAG GAA	GAA GAA GAA	3522
GAG GAG GAG	GAG GAG GAA	GAG GAA	ATG GGA AAC	CCG GAT GGC	3564
TTC TCA CCT	TAG				3576
GCATGCAGGT	ACTGGCTTCA	CTAACCAACC	ATTCCTAACA	TATGCCTGTA	3626
GCTAAGAGCA	TCTTTTTTAAA	AAATATTATT	GGTAAACTAA	ACAATTGTTA	3676
TCTTTTTTACA	TTAATAAGTA	TTAAATTAAT	CCAGTATACA	GTTTTAAGAA	3726
CCCTAAGTTA	AACAGAAGTC	AATGATGTCT	AGATGCCTGT	TCTTTAGATT	3776
GTAGTGAGAC	TACTTACTAC	AGATGAGAAG	TTGTTAGACT	CGGGAGTAGA	3826
GACCAGTAAA	AGATCATGCA	GTGAAATGTG	GCCATGGAAA	TCGCATATTG	3876
TTCTTATAGT	ACCTTTGAGA	CAGCTGATAA	CAGCTGACAA	AAATAAGTGT	3926
TTCAAGAAAG	ATCACACGCC	ATGGTTCACA	TGCAAATTAT	TATTTTGTCTG	3976
TTCTGATTTT	TTTCATTTCT	AGACCTGTGG	TTTTAAAGAG	ATGAAAATCT	4026
CTTAAAATTT	CCTTCATCTT	TAATTTTCCT	TAACTTTAGT	TTTTTTCACT	4076
TAGAATTCAA	TTCAAATTCT	TAATTCAATC	TTAATTTTTTA	GATTTCTTAA	4126
AATGTTTTTT	AAAAAAAATG	CAAATCTCAT	TTTTAAGAGA	TGAAAGCAGA	4176
GTAAGTGGGG	GGCTTAGGGA	ATCTGTAGGG	TTGCGGTATA	GCAATAGGGA	4226
GTTCTGGTCT	CTGAGAAGCA	GTCAGAGAGA	ATGGAAGAAC	AGGCCCTTGC	4276
CAGTAGGTTA	GTGAGGTTGA	TATGATCAGA	TTATGGACAC	TCTCCAAATC	4326
ATAAATACTC	TAACAGCTAA	GGATCTCTGA	GGGAAACACA	ACAGGGAAAT	4376
ATTTTAGTTT	CTCCTTGAGA	AACAATGACA	AGACATAAAA	TTGGCAAGAA	4426

AGTCAGGAGT	GTATTCTAAT	AAGTGTTGCT	TATCTCTTAT	TTTCTTCTAC	4476
AGTTGCAAAG	CCCAGAAGAA	AGAAATGGAC	AGCGGAAGAA	GTGGTTGTTT	4526
TTTTTTCCCC	TTCATTAATT	TTCTAGTTTT	TAGTAATCCA	GAAAATTTGA	4576
TTTTGTTCTA	AAGTTCATTA	TGCAAAGATG	TCACCAACAG	ACTTCTGACT	4626
GCATGGTGAA	CTTTCATATG	ATACATAGGA	TTACACTTGT	ACCTGTTAAA	4676
AATAAAAGTT	TGACTTG CAT	AC			4698

- (2) INFORMATION FOR SEQUENCE ID NO: 6:
(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 25 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein
(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Leu Leu His Arg Tyr Ser Leu Glu Glu Ile Leu Pro Tyr Leu Gly Trp
 5 10 15
Val Phe Ala Val Val Thr Thr Ser Phe
 20

(2) INFORMATION FOR SEQUENCE ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2418 base pairs

(B) TYPE: nucleic acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: genomic DNA

(ix) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GGATCCAGGC	CCTGCCAGGA	AAAATATAAG	GGCCCTGCGT	GAGAACAGAG	50
GGGGTCATCC	ACTGCATGAG	AGTGGGGATG	TCACAGAGTC	CAGCCCACCC	100
TCCTGGTAGC	ACTGAGAAGC	CAGGGCTGTG	CTTGCGGTCT	GCACCCTGAG	150
GGCCCGTGGA	TTCCTCTTCC	TGGAGCTCCA	GGAACCAGGC	AGTGAGGCCT	200
TGGTCTGAGA	CAGTATCCTC	AGGTCACAGA	GCAGAGGATG	CACAGGGTGT	250
GCCAGCAGTG	AATGTTTGCC	CTGAATGCAC	ACCAAGGGCC	CCACCTGCCA	300
CAGGACACAT	AGGACTCCAC	AGAGTCTGGC	CTCACCTCCC	TACTGTCACT	350
CCTGTAGAAT	CGACCTCTGC	TGGCCGGCTG	TACCCTGAGT	ACCCTCTCAC	400
TTCTCCTTC	AGGTTTTTCAG	GGGACAGGCC	AACCCAGAGG	ACAGGATTCC	450
CTGGAGGCCA	CAGAGGAGCA	CCAAGGAGAA	GATCTGTAAG	TAGGCCTTTG	500
TTAGAGTCTC	CAAGGTTTCAG	TTCTCAGCTG	AGGCCTCTCA	CACACTCCCT	550
CTCTCCCCAG	GCCTGTGGGT	CTTCATTGCC	CAGCTCCTGC	CCACACTCCT	600
GCCTGCTGCC	CTGACGAGAG	TCATCATGTC	TCTTGAGCAG	AGGAGTCTGC	650
ACTGCAAGCC	TGAGGAAGCC	CTTGAGGGCCC	AACAAGAGGC	CCTGGGCCTG	700
GTGTGTGTGC	AGGCTGCCAC	CTCCTCCTCC	TCTCCTCTGG	TCCTGGGCAC	750
CCTGGAGGAG	GTGCCCCTG	CTGGGTCAAC	AGATCCTCCC	CAGAGTCCTC	800
AGGGAGCCTC	CGCCTTTCCC	ACTACCATCA	ACTTCACTCG	ACAGAGGCAA	850
CCCAGTGAGG	GTTCCAGCAG	CCGTGAAGAG	GAGGGGCCAA	GCACCTCTTG	900
TATCCTGGAG	TCCTTGTTCC	GAGCAGTAAT	CACTAAGAAG	GTGGCTGATT	950
TGGTTGGTTT	TCTGCTCCTC	AAATATCGAG	CCAGGGAGCC	AGTCACAAAG	1000
GCAGAAATGC	TGGAGAGTGT	CATCAAAAAT	TACAAGCACT	GTTTTCTCTGA	1050
GATCTTCGGC	AAAGCCTCTG	AGTCCTTGCA	GCTGGTCTTT	GGCATTGACG	1100
TGAAGGAAGC	AGACCCACAC	GGCCACTCCT	ATGTCCTTGT	CACCTGCCTA	1150
GGTCTCTCCT	ATGATGGCCT	GCTGGGTGAT	AATCAGATCA	TGCCCCAAGAC	1200
AGGCTTCCTG	ATAATTGTCC	TGGTCATGAT	TGCAATGGAG	GGCGGCCATG	1250
CTCCTGAGGA	GGAAATCTGG	GAGGAGCTGA	GTGTGATGGA	GGTGTATGAT	1300
GGGAGGGAGC	ACAGTGCCTA	TGGGGAGCCC	AGGAAGCTGC	TCACCCAAGA	1350
TTTGGTGCAG	GAAAAGTACC	TGGAGTACGG	CAGGTGCCGG	ACAGTGATCC	1400
CGCACGCTAT	GAGTTCCTGT	GGGGTCCAAG	GGCCCTCGCT	GAAACCAGCT	1450
ATGTGAAAGT	CCTTGAGTAT	GTGATCAAGG	TCAGTGCAAG	AGTTCGCTTT	1500
TTCTTCCCAT	CCCTGCGTGA	AGCAGCTTTG	AGAGAGGAGG	AAGAGGGAGT	1550
CTGAGCATGA	GTTGCAGCCA	AGGCCAGTGG	GAGGGGGACT	GGGCCAGTGC	1600
ACCTTCCAGG	GCCGCGTCCA	GCAGCTTCCC	CTGCCTCGTG	TGACATGAGG	1650
CCCATTCTTC	ACTCTGAAGA	GAGCGGTCAG	TGTTCTCAGT	AGTAGGTTTC	1700
TGTTCTATTG	GGTGACTTGG	AGATTTATCT	TTGTTCTCTT	TTGGAATTGT	1750
TCAAATGTTT	TTTTTTAAGG	GATGGTTGAA	TGAACTTCAG	CATCCAAGTT	1800
TATGAATGAC	AGCAGTCACA	CAGTTCCTGTG	TATATAGTTT	AAGGGTAAGA	1850
GTCTTGTGTT	TTATTTCAGAT	TGGGAAATCC	ATTCTATTTT	GTGAATTGGG	1900
ATAATAACAG	CAGTGGAATA	AGTACTTAGA	AATGTGAAAA	ATGAGCAGTA	1950
AAATAGATGA	GATAAAGAAC	TAAAGAAATT	AAGAGATAGT	CAATTCTTGC	2000
CTTATACCTC	AGTCTATTCT	GTAAAATTTT	TAAAGATATA	TGCATACCTG	2050

GATTCCTTG	GCTTCTTTGA	GAATGTAAGA	GAAATTAAAT	CTGAATAAAG	2100
AATTCTTCCT	GTTCACTGGC	TCTTTTCTTC	TCCATGCACT	GAGCATCTGC	2150
TTTTTGAAG	GCCCTGGGTT	AGTAGTGGAG	ATGCTAAGGT	AAGCCAGACT	2200
CATACCCACC	CATAGGGTCG	TAGAGTCTAG	GAGCTGCAGT	CACGTAATCG	2250
AGGTGGCAAG	ATGTCCTCTA	AAGATGTAGG	GAAAAGTGAG	AGAGGGGTGA	2300
GGGTGTGGGG	CTCCGGGTGA	GAGTGGTGGA	GTGTCAATGC	CCTGAGCTGG	2350
GGCATTTTGG	GCTTTGGGAA	ACTGCAGTTC	CTTCTGGGGG	AGCTGATTGT	2400
AATGATCTTG	GGTGGATCC				2418

Felfe and Lynch
805 Third Avenue
New York, NY 10022

Merry Boon
07/728,838
July 9, 1991



**NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS
CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE
DISCLOSURES**

Mailed: 7/30/91

This application contains sequence disclosures that are encompassed by the definitions for nucleotide and/or amino acid sequences set forth in 37 CFR § 1.821(a)(1) and (a)(2). However, this application fails to comply with one or more of the requirements of 37 CFR §§ 1.821 through 1.825 as follows:

- ☒ 1. This application clearly fails to comply with the collective requirements of §§ 1.821 through 1.825. Applicant's attention is directed to these regulations, a copy of which is attached. *Application lacks a properly formatted sequence listing and a copy of the listing in computer readable form.*
- ☐ 2. This application does not conform exclusively to the requirements of §§ 1.821 through 1.825. The non-conforming material should be deleted. § 1.821(b).
- ☐ 3. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing." § 1.821(c).
- ☐ 4. This application does contain, as a separate part of the disclosure on paper copy, a "Sequence Listing." However, the "Sequence Listing" does not comply with the requirements of §§ 1.821 through 1.825 as follows:
 - ☐ a. The sequence data does not comply with the symbol and format requirements of paragraphs (b) through (p) of § 1.822. Specifically: _____
 - ☐ b. The "Sequence Listing" does not comply with the location and page requirements of paragraph (a) of § 1.823.
 - ☐ c. The "Sequence Listing" does not comply with the information requirements of paragraph (b) of § 1.823. Specifically: _____
 - ☐ d. Other: _____
- ☐ 5. The description and/or claims of the patent application mention a sequence that is set forth in the "Sequence Listing" but reference is not properly made to the sequence by use of a sequence identifier as required by § 1.821(d).
- ☐ 6. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by § 1.821(e).
- ☐ 7. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the computer readable form does not comply with the requirements of § 1.824. Specifically: _____
- ☒ 8. A statement that the content of the paper and computer readable copies are the same has not been submitted as required by § 1.821(f).
- ☐ 9. The amendment to or replacement of the paper and/or computer readable copies of the "Sequence Listing" does not comply with the requirements of § 1.825(a) through (c).
- ☐ 10. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable. Applicant must provide a substitute copy of the data in computer readable form accompanied by a statement that the substitute data is identical to that originally filed. § 1.825(d). Specifically: _____
- ☐ 11. Other: _____

APPLICANT IS GIVEN ONE MONTH FROM THE DATE OF THIS LETTER WITHIN WHICH TO COMPLY WITH THE ABOVE REQUIREMENTS. Failure to comply with the above requirements will result in ABANDONMENT of the application under 37 CFR 1.821(g). Extensions of time may be obtained by filing a petition accompanied by the extension fee under the provisions of 37 CFR § 1.136. Direct the response to, and any questions about, this notice to the undersigned. A copy of this notice MUST be returned with your response.

☒ C. Barnes NT
For: Manager, Application Processing Division

☐ _____
Examining Group _____